ERDC/CRREL TR-00-1

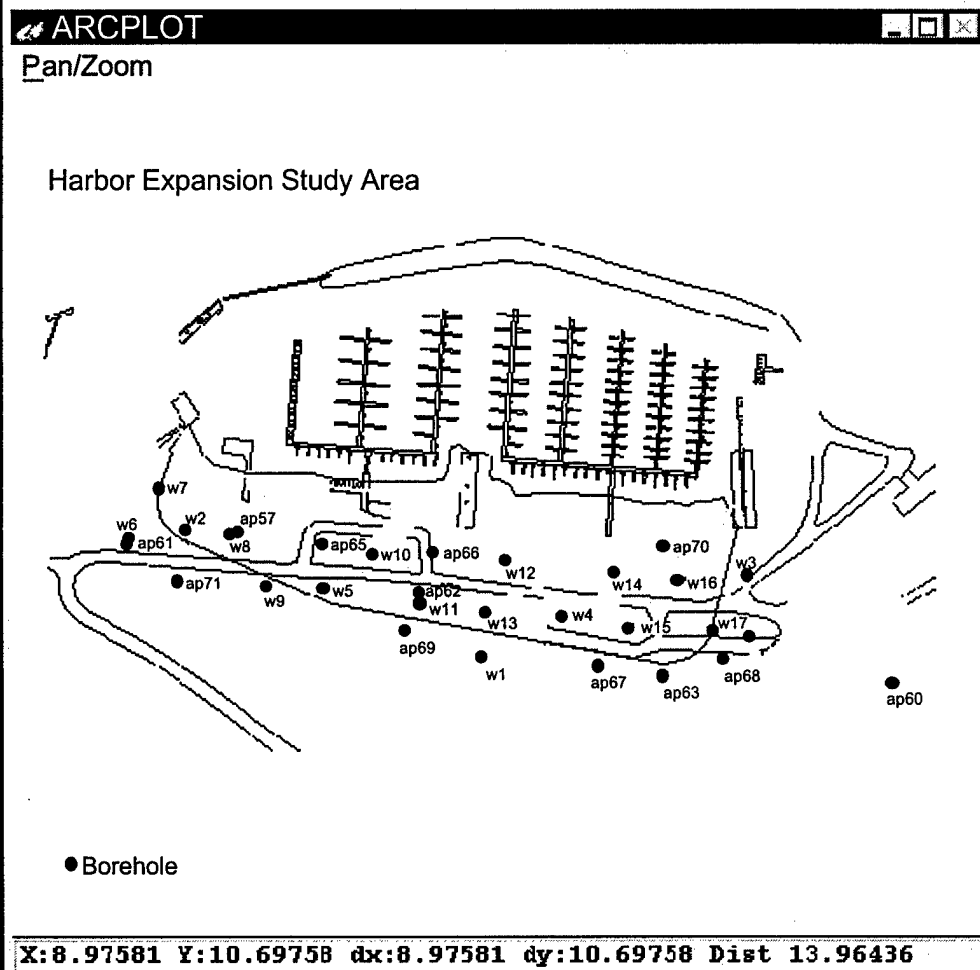Cold Regions Research and Engineering Laboratory

US Army Corps of Engineers®
Engineer Research and Development Center

Integrating CAD Data with Geographic Information Systems Using AutoCAD and ARC/INFO Software

Paul T. Cedfeldt and Mark A. Scott

March 2000

ARCPLOT

Pan/Zoom

Harbor Expansion Study Area

● Borehole

X:8.97581 Y:10.69758 dx:8.97581 dy:10.69758 Dist 13.96436

DTIC QUALITY INSPECTED 2

20000509 102

Technical Report
ERDC/CRREL TR-00-1

**US Army Corps
of Engineers®**
**Cold Regions Research &
Engineering Laboratory**

Integrating CAD Data with Geographic
Information Systems Using AutoCAD
and ARC/INFO Software

Paul T. Cedfeldt and Mark A. Scott   March 2000

PREFACE

This report was written by Paul T. Cedfeldt, Physical Scientist, Geochemical Science Division, U.S. Army Cold Regions Research and Engineering Laboratory, Engineer Research and Development Center, and Mark A. Scott of the Environmental Systems Research Institute, Redlands, California.

Funding for this project was provided by the U.S. Army Corps of Engineers Alaska District.

The contents of this report are not to be used for advertising or promotional purposes. Citation of brand names does not constitute an official endorsement or approval of the use of such commercial products.

CONTENTS

ILLUSTRATIONS

Figure

TABLES

Table

Integrating CAD Data with Geographic Information Systems Using AutoCAD and ARC/INFO Software

PAUL T. CEDFELDT AND MARK A. SCOTT

INTRODUCTION

This document offers the reader a methodology for effectively integrating data created using AutoDesk's computer-aided design (CAD) software AutoCAD (R14) with ESRI's (Environmental Systems Research Institute's) ARC/INFO (version 7.2) geographic information system (GIS) software. Although we intend to educate CAD and GIS users to ways of facilitating the creation of data in CAD for use in GIS, it is not within the scope of this document to consider compatibility issues between every object type in each system. We do, however, discuss a set of issues that are representative of the data integration hurdles faced by many CAD and GIS users.

At the writing of this paper, AutoCAD 2000 had been released but had not yet been tested for this document. ESRI plans to fully support AutoCAD 2000 in upcoming versions of its software. ArcView 3.2 is in production to support AutoCAD 2000 data, and ARC/INFO 8.0 is in beta testing. ARC/INFO 8.0, with its GeoDatabase model, is a new development in ESRI's GIS model. The appendixes and a future revision to this document will include discussion of ArcView 3.1 and ARC/INFO 8.0 in varying levels of detail.

CAD and GIS professionals such as engineers, surveyors, planners, and cartographers commonly integrate AutoCAD drawing (.DWG) or drawing interchange format (.DXF) files, developed internally or by outside contractors, with ESRI formats such as ArcView shapefiles (.SHP) or ARC/INFO coverages. This document assumes that the reader has a professional level of knowledge of AutoCAD or ARC/INFO or both, but we provide base-level information for the sake of completeness.

Although this document is intended for an audience of AutoCAD users, MicroStation users can map their objects to AutoCAD objects, and the concepts should carry over accurately. For example, *layers* in AutoCAD correspond to *levels* in MicroStation, AutoCAD *polylines* correspond to *linestrings* in MicroStation, and *blocks* map to *cells*.

SOFTWARE OVERVIEW

CAD and GIS software, while different, are often used in conjunction with one another. The two packages are concerned with three types of data: coordinate data, descriptive attributes, and graphic symbology. Traditionally, CAD software stores graphic symbology and geometric coordinates in the object information in the drawing database. The final product is often a paper map, so symbology and appearance take precedence during the data creation process. GIS software, however, is primarily concerned with the geographic coordinates (not to be confused with geometric coordinates) and descriptive attributes stored in a relational database. The GIS database is targeted at providing a continuous digital model, connected to a database of attributes, and it is intended for complex query and analysis. Symbology is applied map by map, based on the values in the database.

AutoCAD

AutoCAD is the most widely used PC-based CAD package on the market. It is primarily used in the design and drafting of models in engineering, architecture, and other disciplines, but it also serves as an excellent generic geometric design tool. An engineer might use AutoCAD to design a site for a new mall, and an architect may use it to design a new building. Whatever the discipline, the strength of AutoCAD lies

in its ability to allow the user to manipulate graphic objects quickly and easily to create a finished hardcopy product. Traditionally, AutoCAD is concerned with two types of data: each object in an AutoCAD model stores its location (geometry) and the characteristics of its appearance (graphics). A line may be drawn in a specific color or using a specific line style (dashed, dotted, etc.) that conveys information about its use. Furthermore, AutoCAD uses an organizational approach in its drawing (.DWG) files via a system of layers. These are places where objects can be stored, with the user then being able to manipulate their display as a group. Layers can be assigned colors, although individual objects can have their own color. AutoCAD, being primarily a graphic system, uses the graphic characteristics of an object to carry information about itself. Primitive methods exist in AutoCAD to associate data with drawing entities, such as blocks and object data. One method is to simply create text in the drawing that visually describes an entity but is not rigidly associated with the drawing object. Thus, the CAD package uses a graphic database for symbology and a geometric database for the coordinates of the dimensions of the objects. AutoCAD stores true geometry for circles, arcs, ellipses, and other types of mathematical curves.

ARC/INFO

ARC/INFO is the most widely used professional GIS software package. Professionals in many disciplines use ARC/INFO to analyze relationships among the geographic locations of features and the information describing these features. A GIS is a database in the fullest sense, relating tables of information that describe features to point locations, such as monitoring wells; linear features, such as rivers or roads; and area features, such as county boundaries, parcels of land, or zip codes. The databases, spatial and attribute, are interrelated and can be used to create ad hoc symbology to represent the state of the data, trends in the values, or specific values searched on or queried for. ARC/INFO stores these data in a series of georelational units called *coverages*. The coverages store the coordinates that make up the geometry of the features as well as the database that describes them. Coverages store this information as points (single coordinate locations), arcs (linear connected segments), and polygons (closed areas), and more complex features such as regions.

DATA STORAGE

AutoCAD R14 drawings

AutoCAD R14 stores its data into files called *drawings*. A drawing can contain any design of graphic objects that the user wishes to create, in any organizational mode desired. As we will see, this is both a strength and a weakness when it comes to integrating data. The user can organize data into logical layers and impose a methodology for the symbolization if desired, but this is not required. Although there is a great deal of control over the geometry of the graphics, there are few built-in methods to attach descriptive information via a database. This makes AutoCAD an attractive tool to use for the capture of data because of its powerful editing engine. The task now becomes translating that work into a format that can be converted easily into ARC/INFO to facilitate advanced GIS data manipulation. The coordinate space upon which an AutoCAD drawing exists is a double-precision, 3-D system, originating at 0,0, and extending out in all directions.

ARC/INFO 7.2 coverages

As mentioned, ARC/INFO stores data into logical units called coverages. A coverage stores a set of vector-based geographic features, such as roads, soil boundaries, zip codes, buildings, city locations, sample wells, rivers, etc. Coverages are stored in a georelational format, which means the vector features are associated with a database containing descriptive attributes of the features such as zip code number, soil type, or stream name. A coverage, taken as a whole, is a directory of files, all related together into a model that supports such topological relationships that may need to be created to support network tracing, polygon overlay, and other spatial operations. This topology is created by ARC/INFO's BUILD and CLEAN commands and generates tables to describe the relationships between coordinates that define linear features (arc–node topology) and tables to describe the relationships between linear features that define closed polygonal areas (polygon–arc topology). These relationships create a seamless, nonredundant database that describes the geometry of the geographic data. Another key component of a coverage is that it has a specific coordinate system associated with it, which can be described by its projection and datum. The datum is the ellipsoid upon which the surface of the earth is approximated, and the projection is the mathematical method via which the surface is "flattened" into a Cartesian coordinate mapping space. Figure 1 illustrates data storage concepts for both AutoCAD and ARC/INFO.

ARC/INFO coverage features

ARC/INFO is capable of storing data in many simple and complex features. Some examples of complex feature classes are routes (paths through multiple arcs), regions (collections of polygons in potentially overlapping areas), and raster data sets such as images and grids. We will concern ourselves here with the simple 0-D,
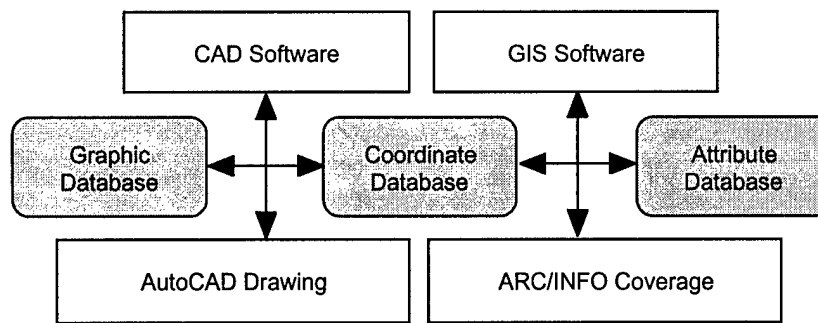
Figure 1. CAD and GIS software data storage.

1-D, and 2-D features of points, arcs, and polygons. These will then map through the conversion from AutoCAD points, block inserts, lines, polylines, light-weight polylines, circles, arcs, and other drawing enti-ties. Each ARC/INFO feature is connected to its fea-ture attribute table via an internal ID system. This is not a major concern to the AutoCAD user. However, creating a way to uniquely identify each entity by a system-dependent identifier might be a concern. For example, each fire hydrant in a drawing may have some unique number assigned by the Public Works Depart-ment. This ID might be placed in the drawing to iden-tify the hydrant. A CAD user might be satisfied to place this on the drawing to be plotted out. The GIS would need to store this value with the hydrant in the feature attribute table, so it can join it to other databases in the future. As we will see, there are various methods to capture this type of information into ARC/INFO.

Point features

Points represent geographic features that have no area or length, such as sewer manholes, bank locations, accident locations, or features that are too small at a given scale to be represented by their dimensions. Points are stored as single X,Y coordinate pairs. In a GIS, descriptive data are stored in a database, then related to the spatial database that holds the coordinates. As shown in Figure 2, the spatial database stores the coordinates

in a LAB file that is linked internally to the point attribute table (PAT).

Arc features

Arcs represent both linear features and the borders of polygon features. Linear features represented by arcs can have length but no area, such as a contour line, or can be long narrow features whose width is not appar-ent at a given map scale, such as a river. Each linear feature may be made up of many arcs. Since arcs also form the basic linear infrastructure from which com-plex GIS processes are derived, such as tracing and surface computation, it is very important that an entity be continuous. An arc is a continuous string of X,Y coordinate pairs (vertices), beginning at one location and ending at another, stored in an ARC file. Geometri-cally, an arc is similar to an AutoCAD Polyline entity, which is often used to create ARC/INFO arc features. The topological structure is created so that the starting and ending points are special locations known as *nodes*. The in-between points are the vertices that define the arc's shape. This structure is used to trace linear sys-tems and to form the boundary of polygon features. This means that there should be no break in the continuity of entities drawn in AutoCAD to represent features in ARC/INFO. If a text entity is used to indicate the diameter of a water line, it should not break the conti-nuity of the water line and should be placed above or
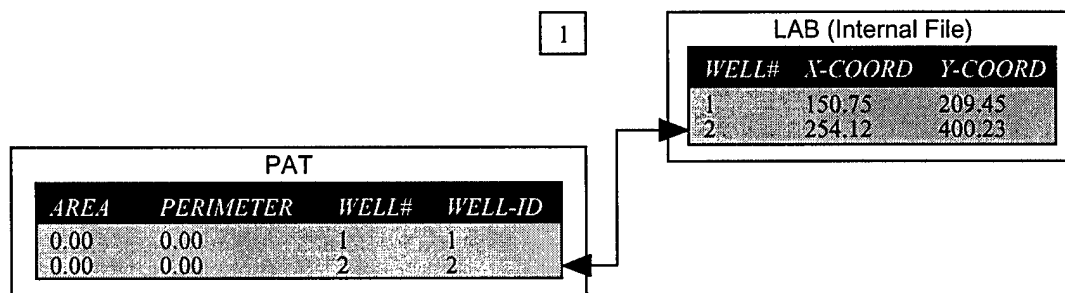

Figure 2. ARC/INFO point feature.

3

below the line. The AutoCAD user must, in this situation, remember that the final destination of these mapped data is not a paper plan, used for visual purposes, but a digital model to be used for analysis. Other tools in ARC/INFO can be used to manipulate the location of descriptive text. The arc attribute table (AAT) and node attribute table (NAT) are used to store the attributes (Figure 3).

Polygon features

Polygons are used to represent area features, such as states, counties, lakes, and land-use zones. Polygons enclose areas that meet a user-specified set of common characteristics for the phenomena being represented. In the ARC/INFO model, a series of arc features is used to create a closed area that defines the polygon feature. A polygon–arc–list (PAL) file defines the relationship. Figure 4 illustrates this relationship graphically. The attributes are stored in the polygon attribute table (PAT). An additional characteristic of a polygon feature is that it has a special point that lies within its border called the *label point*. A label point is used by ARC/INFO to identify the polygon, because the polygon is actually not the arcs that make up its boundary but a set of relationships between those arcs. Any AutoCAD entity that can be used to represent a point feature can be used to model the label point of a polygon. Polygons can also model holes or voids in a region. When creating data to be converted into ARC/INFO polygons, CAD users should snap end points of lines together and not create regions where the areas of polygons would overlap. The

ARC/INFO CLEAN command can be used to repair overlapping and unconnected areas, but it is preferable to have the data created appropriately at the time they are digitized. Polygon features will not be directly created at the conversion level but will be assembled from the geometry that defines the arc features after using CLEAN.

Annotation features

Annotation features are used to store descriptive text that may or may not be connected to spatial features, such as points, arcs, and polygons, in the coverage. Annotation may be derived from the database values of a point, line, or polygon feature, or it can be standalone. Annotation can be captured from AutoCAD and stored in the coverage for later use, mostly for creation of cartographic map output.

MODELING ARC/INFO FEATURES WITH AUTOCAD ENTITIES

This section focuses on how to represent information in an AutoCAD drawing with AutoCAD entities, with the purpose of converting them into ARC/INFO features.

Point features

Discrete coordinate locations can be represented in the drawing in a number of ways, all of which can be converted into ARC/INFO point features. The AutoCAD point entity is a location defined in the draw-
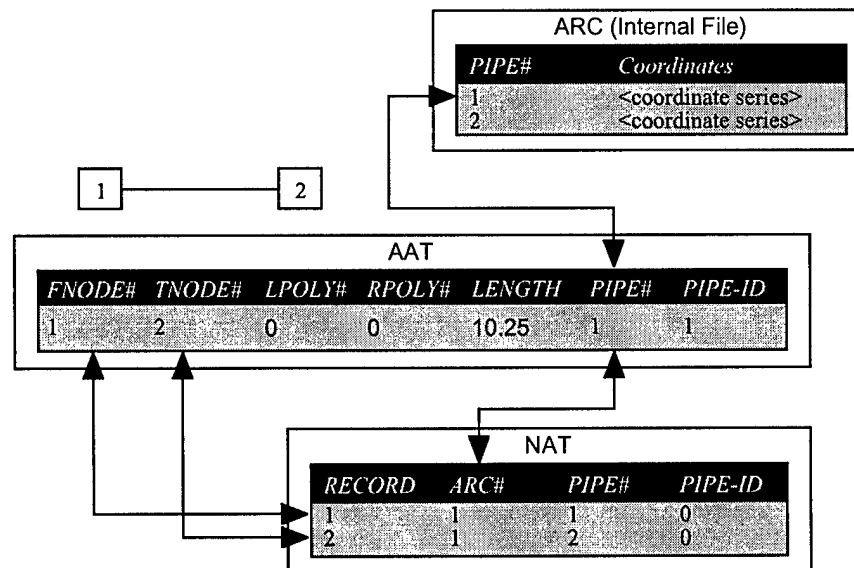

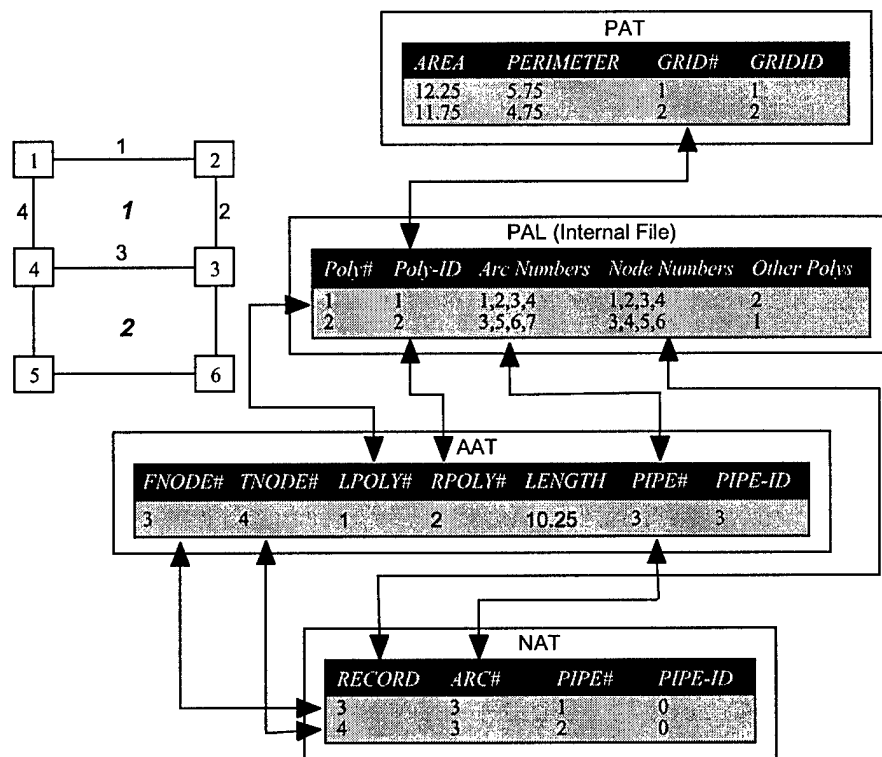**Figure 3. ARC/INFO arc/node features.**

Figure 4. ARC/INFO polygon features.

ing that can have several shapes. The AutoCAD shape entity has its definition stored in a shape font definition file. Either of these can be converted into point features. The AutoCAD block entity can be used to capture a multitude of data types. A *block* is a compound object, defined from a set of other AutoCAD entities. For example, a fire hydrant symbol could be created from a set of lines, arcs, and polylines, drawn in various colors and on various layers, to look like a fire hydrant. The block is then defined and saved as a template. This block can be placed multiple times in the drawing. Each individual instance of the block is called an *insert*. Each block insert can be captured as a point, based on the location where it was inserted into the drawing. There are options to create ARC/INFO arc features from the individual geometric parts of the block, but these must be valid object types. In addition to the symbol, a block can contain a set of *attributes*, which are items of descriptive text that can accompany the symbol. The block definition contains a set of attribute definitions called *tags*, and each instance of the block (insert) contains the tag attributes. Our hypothetical hydrant block could have two attributes associated with it: the hydrant ID number and the date of installation. These can be captured as text annotation, but they can also be stored as a point feature.

Arc features

AutoCAD line objects are simple two-point objects that delineate an edge of a 2-D or 3-D entity. These objects translate directly into ARC/INFO arc features. The problem from a GIS perspective is that any break in the line results in a break in the connectivity of the feature. Lines that are broken in two, for example, to place a piece of text, can result in difficulty in GIS model creation. Several line entities can be strung together to form linear boundaries, but they should be snapped so the end points coincide. A more appropriate AutoCAD object to use would be the polyline or LWpolyline (light-weight polyline) object. These are multivertex linear objects that can be used to symbolize a road centerline, river edge, or any other 2-D object. LWpolylines are often used when there is no need for 3-D elevations on the vertices, since they take up much less storage in the AutoCAD database. Each segment of a polyline can contain a bulge factor, making its segment a geometric arc. Again, these entities should be continuous if they are being prepared as layers to be converted into ARC/INFO coverage arc features.

AutoCAD geometric arcs and circles can be converted to ARC/INFO arc features and will be tessellated into segments approximating the curved segment, with one segment per degree of arc. The primary entity

5

used in AutoCAD to digitize linear features is the polyline. An AutoCAD polyline is a multivertex, multisegment, and two-dimensional entity. Three-dimensional polylines can be created in AutoCAD, but those with vertices at varying Z coordinates should be avoided, as the conversion process will ignore them. If the polyline is at a consistent elevation, it can be captured for future use, as we will see. As noted, the lightweight polyline, which is a simple, noncurved, 2-D version of a polyline, may be used. While AutoCAD users often prefer these entities because of their low overhead, it does not matter for the conversion process, and they are treated as a regular polyline. Any further reference to an AutoCAD polyline will imply a polyline or lightweight polyline object. Note that ARC/INFO will create *nodes* at the end points of the line, polyline, or arc entities when they are converted to ARC/INFO arc features. Figure 5 illustrates ARC/INFO arc features and corresponding AutoCAD entities.

Polygon features

As noted previously, an ARC/INFO polygon feature is created from arc features via the CLEAN command. This being said, information that is intended for conversion into ARC/INFO arc features, then CLEANed into polygon features, should be carefully digitized to avoid error. Polylines could be digitized either as closed (an option when creating the polyline), or open, with the start and end points identical. If these methods are used, closed and snapped polylines will be easily converted into ARC/INFO polygon features. Digitizing polygons that overlap will result in the creation of multiple polygons, as this is not allowed in the polygon model. This may also lead to undesirable results. If the line work is digitized and snapped correctly, CLEAN will create polygons and drop any redundant geometry. This means that a closed or open polyline that dupli-

cates an edge is not necessary, because of the nature of the coverage model and the CLEAN process.

Annotation features

Text objects are used to indicate additional attributes, such as pipe conditions or installation dates, or simple labels on objects, such as a river name. This text, however, has no connection to the object, rather it simply provides a visual aid on the drawing canvas. The text object represents data that can be captured from the drawing as ARC/INFO annotation. Each piece of text can be captured, along with its size, value, and rotation angle, as annotation or as a point feature located at the insertion point of the text. No information concerning the font or AutoCAD text style is needed. While it is useful for the data provider to deliver necessary text fonts to the ARC/INFO user, it is not crucial to the conversion process because ARC/INFO uses its own set of text symbols.

Attributes, as noted in the section on point features above, are textual objects that are part of a compound block insert. This text, in addition to being captured as point data and linked back to the symbol, can be captured as annotation and treated as such in the GIS database. These data will follow the same rules and procedures as text.

AUTOCAD SYMBOLOGY AND DATA ORGANIZATION

AutoCAD organizes its data, stored in file-based drawings, into a series of overlays called *layers*. In addition to layers, the user can utilize methods of symbolization such as block symbols, color, line type, and line width to visually describe the drawing entities. Attribute information can be carried along in the form of AutoCAD block attributes as well.
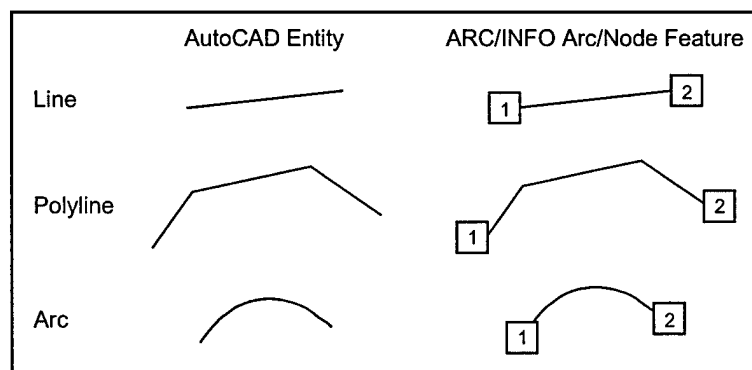

	AutoCAD Entity	ARC/INFO Arc/Node Feature
Line		
Polyline		
Arc		

Figure 5. Representing ARC/INFO arc features with AutoCAD entities.

Layers

A layer is a free-form storage entity capable of holding data in various symbolic patterns and containing data of various object types. Although there is no requirement to do so, users often organize their layers into different types of information. This organization is commonly grouped by some scheme meant to lend structure to the drawing. For example, a drawing could be broken up into CAD layers called parcels, streams, streets, and hydrants. The parcels layer could contain the polylines that make up the boundary of the parcels, as well as the text that describes the parcels. The streams layer might contain the lines delineating the streams, as well as the stream names. The hydrant layer could contain AutoCAD block symbols that have a hydrant symbol, as well as some block attributes describing the hydrant. Layers can be named using alphanumeric names. An effective layering scheme is essential for the conversion process. When possible, all text objects should be on layers independent of the entities they describe. For example, street centerline entities could be placed on a layer called GIS_STREET_LINES, and text labeling for these streets would be stored on a layer called GIS_STREET_TEXT. Using separate layers for data that will be stored in different feature classes in the coverage is very helpful in the conversion process. All non-GIS data should be isolated on layers that can be easily identified and excluded from the conversion process. All GIS layers should contain only one feature class, such as POINT, LINE, or TEXT, again to facilitate conversion.

Colors

Proper use of color can convey much information about the data in a drawing. Colors can be assigned on a layer-by-layer basis (BYLAYER), meaning that the layer has a default color and all objects drawn on it are of that color, and they can be assigned on an object-by-object basis (BYCOLOR). Colors can be used to further refine the definition of objects on a layer. A layer called WATER_LINES could contain the water pipe information on a site plan. The pipe material could then be symbolized by colors on this layer; for example, red lines could represent PVC, blue lines cast iron, and green lines ductile iron. This provides a way for the AutoCAD user to carry pseudo-attributes on each object without having to design object data tables with another application (such as AutoCAD Map) or writing a custom application. Often text objects are used to indicate additional attributes, such as pipe conditions or installation dates or both. This text, however, has no connection to the water pipe object, rather it simply provides a visual aid on the drawing canvas.

Line types

LineType libraries are used in AutoCAD to indicate varying attributes. For example, a dashed line might be used to indicate an intermittent stream. This method of showing breaks in a linear feature is acceptable for translation; however, physically "breaking" the line to place text or another symbol there is not acceptable, because the continuous nature of the linear feature must be preserved.

Line width

In reference to the example in the Layers section above, water lines could be drawn using specific line widths to represent the pipe diameter: a line with a width of 4 could mean a 4-in.-diameter pipe. This methodology can break down rather quickly when there is a large amount of data or a diverse number of diameter values, such as 42 values requiring 42 different line widths.

PROBLEMS TO AVOID

There are certain entities that should never be used for representing CAD data to be converted using ARC/INFO. They can be used, but they need to be isolated on non-GIS layers.

Using 3-D data

Three-dimensional data can be used on GIS layers, but there are some limitations. Polylines, or any other object that has multiple parts on different elevations, will be reduced to the elevation of the first entity. Contour lines and other entities that are at a single elevation (Z coordinate) can be converted properly.

Model space and paper space

Paper space is an environment in the AutoCAD drawing in which multiple viewports can be created to contain actual drawing data that has been rotated, clipped, and annotated for use in plotting to hard copy. Paper space is, by definition, in the coordinate space of the plotted output and can have no connection to the geographic coordinate system. While these data are useful in CAD, paper space objects and viewports are invalid for use in ARC/INFO and cannot be converted. Therefore, *model space* should always be used.

Details and title blocks

Any details, insets, and title block information pertinent to the drafting of the drawing should be isolated onto non-GIS layers.

Groups

An AutoCAD *group* is a collection of heterogeneous objects—such as a street line, its name text, and two

sign symbols—organized into one object. This is similar in structure to a block, but it cannot be converted into GIS data and should be avoided.

Multilines and other entities

A *multiline* is a special type of line that can be defined as sets and patterns of parallel lines treated as a single line. These lines cannot be converted using the DXFARC conversion command in ARC/INFO. Rays and Xlines are construction objects in CAD and should never be used to symbolize GIS information.

Hatching

An AutoCAD hatch object is a specially defined pattern used to fill in an enclosed area. The hatch is created as an anonymous (unnamed) block in the database placed at coordinate 0,0,0. Hatch patterns cannot be converted and should not be used to convey any information about an enclosed area.

External references

An AutoCAD external reference (XREF) is a reference to another AutoCAD drawing, basically placing its contents into another drawing as a block. XREF objects cannot be interpreted by ARC/INFO and should not be used.

Dimensions and leaders

A *dimension* (DIM) is a complex object consisting of text, lines, and other symbols used to indicate sizes, distances, and other quantities. DIMs cannot be converted and should not be used. Simple text objects should be used instead. A *leader* is an entity that is used to draw an arrow to point to an object. It should be treated as a dimension and similarly avoided.

Blocks as grouped graphics

Blocks are very useful for symbology and locating point features, but a block entity should not be used when it is created as a large-scale collection of graphics, say, from another drawing. If the CAD designer inserts a large portion of another drawing into the existing one, he can insert it with the explode option, which will break up the graphic elements individually. This is always the necessary procedure. If blocks are used as symbols, the BLOCKS subcommand option of the DXFARC command can be used.

Layer 0

Layer 0 (zero) is a special layer in an AutoCAD drawing, and it is best not used to draw any objects to be converted via DXFARC. Blocks should be defined on specific layers, as should their attributes. If all block definitions and attributes are on layer 0, this can make the conversion process quite difficult.

GEOREFERENCING AND COORDINATE SYSTEMS

Several additional points of interest should be noted as they apply to the conversion of CAD data. Most notably, GIS data need to be associated with a specific map projection and datum. For this reason, the CAD data should be digitized in a coordinate system that is agreed to by all parties involved and in an agreed-upon set of units (meters, feet, etc.). If the drawing data are not referenced directly, then there should be at least four control points on the drawing at which the geographic coordinates are known.

METHODOLOGY

Adoption of a standard and agreed-upon method to create GIS data for use in ARC/INFO will make the data provider's job, and the GIS professional's job, much easier and more efficient. In this section, we discuss several concepts that can be very useful.

Layering standards and conventions

The adoption of a consistent and agreed-upon layering standard is critical to the success of the conversion. In this way both the data provider and the GIS user know what to expect in the drawing without having to be experts in both systems. One of the most effective and simplest concepts is to isolate GIS and non-GIS data into their own named sets of layers. For example, all GIS data could be stored on layers beginning with the prefix GIS_. Data for street centerlines could be stored on a layer called GIS_STREET_LINES. All layers that do not start with the GIS_ prefix, although still useful to the CAD drafter, could be ignored by the ARC/INFO user performing the conversion. Another useful standard is to place all feature classes (point, arc, polygon, and annotation) onto separate layers. The layer GIS_STREET_LINES implies that the layer contains only line (arc) data for the street centerlines. Text for the street names could be stored on GIS_STREET_TEXT. Again, this makes it easier to convert the data. Text especially should be placed on an individual layer. It is even useful to place blocks and attributes on individual layers. If there is a layer for each attribute tag in a block, the GIS user can easily extract the pieces needed to complete the conversion.

Color and line type schemes

As mentioned earlier, color can be used to catego-

8

rize data effectively within a layer. Water lines could be drawn in specific colors to indicate material and with varying linewidths to indicate pipe size. A line on the layer GIS_WATER_LINES drawn in blue and with a line width of 3 might indicate a stainless steel pipe with a diameter of 36 in. These color categorizations and line widths would need to be agreed upon beforehand. After that, it is very easy for the ARC/INFO user to capture the information.

Using text or block inserts as data tags

Text can be used to create ARC/INFO annotation features to help describe features. In addition, ARC/INFO point features could be created, then the XCODE file used to get the text value. These values can be used to "tag" the feature. For example, a set of lines indicating a parcel's boundary could have the parcel ID number placed inside the parcel. The text could be captured as a point feature, the text value related via the XCODE file, and then a point-in-polygon overlay operation in ARC/INFO could be used to match them together. The same technique could be employed with linear features by using the ARC/INFO NEAR command to find the closest line feature to a point feature. The same can be said for block inserts, where the XCODE file contains the block attributes, which contain valuable information about the feature.

CONVERTING THE DATA: DXFOUT AND DXFARC

In AutoCAD, the command to export a .DXF file is DXFOUT. The menu selection Save As... can also be used to save the drawing as a .DXF file. ARC/INFO can read .DXF files from AutoCAD releases 12, 13, and 14 in either ASCII or binary format. The ASCII format creates the largest files, but they are easier to examine for problems if the user does not have the AutoCAD software. DXFOUT can be used to output specifically selected objects or the entire drawing. If desired, the AutoCAD user could export specific layers using the object filtering capabilities. This could help keep the size of the .DXF files reasonable. Either way, ARC/INFO can integrate the layers from the entire file or from subsets of the file.

The DXFARC command in ARC/INFO is the primary tool via which AutoCAD data, saved as a drawing interchange format (.DXF) file, are converted into ARC/INFO coverages. DXFARC can selectively process layers from the .DXF file. For a complete description of the use of the DXFARC command, refer to the ARC/INFO documentation or on-line help. The rest of this section borrows from that help document but does not cover all options.

There are several components to the use of this command, but the most important are the subcommand options and the INFO tables that can be used to capture the AutoCAD properties of the data. The following subcommands can be used to process specific parts of the .DXF file:

ALL	Converts all entities except BLOCKS. This option is the default if no option is given. This is the same as specifying ARCS, POINTS, INSERTS, TEXTPOINT, TEXTANNO, ATPOINT, and ATANNO for a single layer.
ARCS	Converts entities to create coverage arcs. DXF entity types that can be converted to coverage arcs include line, 3Dline, trace, solid, 3Dface, circle, arc, polyline, and LWpolyline.
POINTS	Converts point entities (point, shape) from the layer.
TEXTPOINT	Converts text entities (text) from the layer into point features.
TEXTANNO	Converts text entities (text) as annotation in subclass named DXF. Each layer will create a unique annotation level (sequential, in order of appearance). Determine the sequential order by reviewing the output of DXFINFO.
ATPOINT	Converts attribute and ATTDEF from the ENTITIES section of the layer as point features.
ATANNO	Converts attribute and ATTDEF from the ENTITIES section as an annotation in a subclass named DXF. Each layer creates a unique annotation level (sequential, in order of appearance). Determine the sequential order by reviewing the output of DXFINFO.
INSERT	Converts inserts as point features.
BLOCKS	Performs the equivalent of exploding blocks for all points, lines, or multipoint lines. Text entities contained within blocks will be converted if either TEXTANNO or TEXTPOINT is used. Attributes will be converted if either ATANNO or ATPOINT is used. Inserts will not be converted as point features.

When converting data into ARC/INFO arcs, the properties of the linear features are stored in an INFO table with the extension .ACODE. The properties of

data converted into point features are stored in an INFO table with the extension .XCODE.

When run through DXFARC, a coverage called CADDATA would produce an INFO table called CADDATA.ACODE containing the following items:

CADDATA-ID	ID of the coverage feature in its attribute table.
DXF-LAYER	Layer to which the entity belongs.
DXF-COLOR	Color of the entity.
DXF-THICKNESS	Thickness of the entity.
DXF-TYPE	Line type (continuous, dashed, etc.) of the entity.
DXF-ELEVATION	A single elevation value for the entire entity: its Z coordinate.

The CADDATA coverage could also produce an INFO table called CADDATA.XCODE containing the following items for the points:

DXF-LAYER	Same as for ACODE file.
DXF-COLOR	Same as for ACODE file.
DXF-ELEVATION	Same as for ACODE file.
DXF-ANGLE	Rotation angle in degrees.
DXF-TYPE	Line type for the point feature.

If a layer is converted by DXFARC with the TEXTPOINT option, the XCODE file contains these additional items:

DXF-ANGLE	Angle of the text.
DXF-SIZE	Size of the text.
DXF-TEXT	Text value for the DXF-ATTRIB. If the entity is a point, shape, or insert, this field is blank.

If a layer is converted using the ATPOINT option, these additional items are added to the XCODE file:

DXF-ATTRIB	Holds the BLOCK name for IN-SERT points, the shape name for SHAPE entities, and the tag or item name for ATTRIB and ATTDEF entities.
DXF-TEXT	Stores the item value for ATTRIB and ATTDEF entities.
DXF-ATTRIB	Blank for text and point entities.
DXF-IID	Used to link attributes to the corresponding INSERT entity. Each INSERT entity can have multiple attribute entities. IN-SERT entities are converted to point features with a unique user ID. This user ID is stored in a DXF-IID item of all the attribute entities associated with that IN-SERT.

ARC/INFO users can link attribute information from the ATTRIB point features to the point features from the insert entities using the ARC/INFO RELATE command. See the ARC/INFO documentation for more information.

SUMMARY AND PROCESS STEP EXAMPLE

To summarize, Table 1 lists the AutoCAD data types to use in a drawing and the resulting data types after conversion to an ARC/INFO coverage.

The following example is based on an AutoCAD drawing that contains information about the water distribution system on a site (Figure 6). The water lines are stored on an AutoCAD layer called GIS_WATER. The AutoCAD user digitizes these lines using the AutoCAD POLYLINE command. The AutoCAD DXFOUT command is used to create the WATER.DXF file. Using DXFARC, the ARC/INFO user can now extract the arc features and use BUILD. This creates the ARC/INFO arc and node features and their associated attribute tables. The water line information is stored in an AutoCAD block, with three attribute tags to store the pipe length, pipe identifier, and pipe type on an AutoCAD layer called GIS_WATER_DATA. The AutoCAD user inserts these blocks using the AutoCAD INSERT command and fills in the data fields when prompted. The .DXF file is once again exported using DXFOUT.

Using DXFARC, the ARC/INFO user can now extract the point features from the layer and BUILD the point and annotation feature classes. Figure 7 details the command syntax used for the DXFARC command. The .XCODE file contains the actual attribute values for the points created in the DXFARC process. This file can be joined back to the point attribute table (PAT) in ARC/INFO. Features with the same DXF-IID field all belong to the same block insert.

To find the nearest water line to the block insert point features, use the ARC/INFO NEAR command. Figure 8 details the command syntax used for the example. Note in the figure that no polygon features are actually created at the time of conversion, rather, they are created by subsequent use of the ARC/INFO CLEAN command. As discussed above, as long as the geometry of the arcs is created correctly, correct polygon topology will follow.

Figure 9 shows the converted data, and Figure 10 summarizes the entire conversion process as a flow-chart.

Table 1. AutoCAD data types and ARC/INFO coverage equivalents.

Draw in AutoCAD:	To represent in ARC/INFO:
Point	
Insert (block)	Point feature
Attribute	
Shape	
Line	
Polyline	Arc feature
Arc	Polygon feature (after CLEAN)
Circle	
Closed polyline	Polygon feature (after CLEAN)
Closed lightweight polyline	
Text attribute	Annotation feature

Figure 6. AutoCAD drawing containing water data.

```
Arc: dxfarc
Usage: DXFARC <in_dxf_file> <out_cover> { text_width}
{ attrib_width}
Arc: DXFARC WATER.DXF WATER 32 32

Enter layer names and options (type END or $REST when done)
===============================================================
Enter the 1st layer and options : GIS_WATER ARCS
Enter the 2nd layer and options : END
Do you wish to use the above layers and options (Y/N)? Y

Processing D:\WORK\CRREL\COVERS\WATER.DXF ...
No labels, killing XCODE...
Externalling BND and TIC...

        172 Arcs written.
          0 Labels written.
          0 Annotations written.
          0 Annotation levels.
Arc: BUILD WATER LINES
 Building lines...
Arc: RENODE WATER
288 unique nodes built for D:\WORK\CRREL\COVERS\WATER
Arc: DXFARC WATER.DXF WATER_TAGS 32 32

Enter layer names and options (type END or $REST when done)
===============================================================
Enter the 1st layer and options : GIS_WATER_DATA ATPOINT
ATANNO INSERT
Enter the 2nd layer and options : END
Do you wish to use the above layers and options (Y/N)? Y

Processing D:\WORK\CRREL\COVERS\WATER.DXF ...
No arcs, killing ACODE...
Externalling BND and TIC...

          0 Arcs written.
        688 Labels written.
        516 Annotations written.
          1 Annotation levels.
Arc: BUILD WATER_TAGS POINT
 Building points...
Arc: BUILD WATER_TAGS ANNO.DXF
 Building annotation...
Arc:
```

Figure 7. Using the ARC/INFO DXFARC command.

```
Arc: near
Usage: NEAR <in_cover> <near_cover> { LINE | POINT | NODE}
{ search_radius}
           { out_cover} { NOLOCATION | LOCATION}
Arc: NEAR WATER_TAGS WATER LINE 1.0 WATER_TAGS2
Identifying features in WATER near those in WATER_TAGS...
```

Figure 8. Using the ARC/INFO NEAR command.

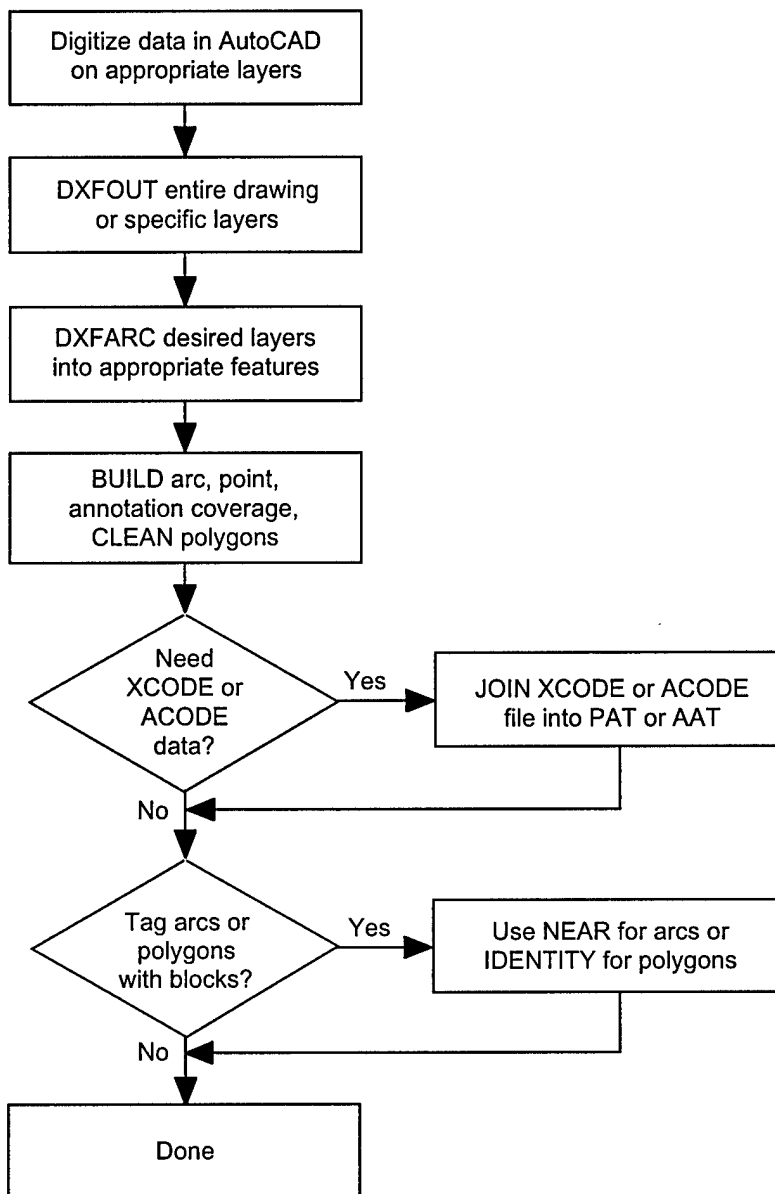**Figure 9. ARCPLOT canvas showing converted data.**

Figure 10. Flowchart showing conversion from AutoCAD to ARC/INFO.

APPENDIX A: ARCVIEW GIS 3.1

ArcView stores data in a format called a *shapefile* (.SHP). The shapefiles are defined, for our purposes, as point, line, or polygon shapes. There is no text or annotation data type in a shapefile. Geometry is stored in a .SHP file, a spatial index is built in a .SHX file, and database attributes associated with the shapes are stored in a dBase file (.DBF). AutoCAD data can be read directly using the ArcView CAD Reader extension. Conversion is facilitated via the Convert To Shapefile option in ArcView. Custom conversion tools can be written in ArcView with the Avenue programming language. See the ArcView help document for object mapping.

ArcView has two additional features that make it useful for the conversion process. First, any AutoCAD polyline that is closed, or any open polyline whose beginning and ending points are identical, is interpreted as a polygon shape. Second, attributes of blocks are automatically associated with the block insertion point, which defines a point shape, by denormalizing the data in the associated table.

A future version of this document will provide more information on using ArcView GIS and CAD data.

APPENDIX B: ARC/INFO 8.0

ArcInfo 8.0 stores data in a GeoDatabase. This GeoDatabase can exist in a standalone format using a Microsoft Access file (.MDB), or it may be stored in a relational database management system (RDBMS) such as Oracle or SQL Server to be used in a multiuser environment. The GeoDatabase stores information in a similar fashion to an ARC/INFO coverage, but all in one unit. Since the object models of AutoCAD and ARC/INFO are visible via Visual Basic or any other Microsoft Windows COM-compliant development environment, creation of custom translators would be very feasible in this powerful environment. See the ArcInfo 8.0 help document for object mapping.

A future version of this document will provide additional information on using ARC/INFO 8.0 and CAD data.

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
March 2000	Technical Report	

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Integrating CAD Data with Geographic Information Systems Using AutoCAD and ARC/INFO Software	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)	5d. PROJECT NUMBER
Paul T. Cedfeldt and Mark A. Scott	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
U.S. Army Engineer Research and Development Center Cold Regions Research and Engineering Laboratory 72 Lyme Road Hanover, New Hampshire 03755-1290	ERDC/CRREL TR-00-1

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR / MONITOR'S ACRONYM(S)
U.S. Army Corp of Engineers Alaska District FUDS Program P.O. Box 898 Anchorage, AK 99506-0898	11. SPONSOR / MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

Available from NTIS, Springfield, Virginia 22161.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Software for computer-aided design (CAD) and geographic information systems (GIS) are often used in conjunction with one another. This document proposes a methodology for effectively integrating data created using Autodesk's AutoCAD R14, the most widely used CAD software package, with Environmental System Research Institute's (ESRI's) ARC/INFO 7.2, the most widely used professional GIS software package. A brief discussion of software data models is provided, followed by an enumeration of certain entities that should never be used to represent AutoCAD data that is going to be converted using ARC/INFO. The document concludes with an example implementation of the proposed integration methodology.

15. SUBJECT TERMS		
ARC/INFO	Computer-aided design (CAD)	Spatial data translation
AutoCAD	Geographic information	

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
U	U	U	U	24	